Social Network Analysis and Mining (2024) 14:10
https://doi.org/10.1007/s13278-023-01175-9

ORIGINAL ARTICLE

Shaping climate change discourse: the nexus between political media landscape and recommendation systems in social networks

Yousef Sakieh[1]

Received: 14 March 2023 / Revised: 2 September 2023 / Accepted: 24 November 2023

Abstract

Recommendation systems have become ubiquitous, and they actively participate in creating our individual and collective identity. In this paper, the diffusion of climate change information has been studied based on YouTube's recommendation system and the political media landscape. The YouTube channels of CNN, BBC News and Fox News, as the most popular channels, respectively, for Left, Center and Right parties, were explored using web scraping and social network analysis to check what kind of recommended content will pop up if a user looks for climate change videos. Using an agent-based modeling approach, the competition between Left, Center and Right media in pushing their own narrative of climate change in society was simulated. The results suggest YouTube's recommendation algorithm is highly biased since most of the recommended content was from the same channel fitting their own political agenda. The agent-based modeling indicates the size of a network is a decisive factor in further spread of a message as Left media always dominated Center and Right media in pushing their own perspective on climate change regardless of higher weights assigned to Right media. This study shed light on how public perception on climate change can be shaped by recommendation systems and digital companies.

Keywords Recommendation systems · Social networks · YouTube · Climate change · CNN · Fox news

1 Introduction

1.1 Contextual framework

Recommendation systems have become ubiquitous. In the age of artificial intelligence and social media, the contemporary humans live in a digital matrix, where recommendation systems have the authority to tell them what movie to watch, what person to meet, what restaurant to go and what news to read and they have been integrated in almost every aspect our lives. Companies and digital businesses rely on recommendation systems to increase their profit, and we as users rely on them as they reduce information overload and make our lives easier. They boost potential online engagement, reduce customers churn, maximize user activity on different platforms and increase sales and market share. Recommendation systems offer mutual benefits to producers and consumers. They optimize micro- and macro-customer targeting and improve key performance indicators on customers. They actively participate in creating our individual and collective tastes, and in the long run, they participate in shaping our identity with important ethical implications (Milano et al. 2020, 2021).

Recommendation algorithms have far-reaching impacts on society, and understanding these effects is crucial for various reasons. Firstly, they can inadvertently contribute to the creation of information bubbles and filter bubbles. These algorithms prioritize content similar to users' previous engagements, potentially reinforcing biases and limiting exposure to diverse viewpoints. Additionally, recommendation algorithms filter out content that does not align with a user's interests, narrowing perspectives and hindering exposure to dissenting viewpoints, which is vital for fostering a more informed and open society.

Secondly, privacy concerns arise as recommendation algorithms rely on user data for personalization. The collection and analysis of personal information can lead to privacy breaches. The study of recommendation algorithms is essential to assess the ethical implications of data collection and usage, ensuring that user privacy is respected. Moreover, these algorithms significantly impact monetization and

✉ Yousef Sakieh
yousef.sakieh@helsinki.fi

[1] Faculty of Biological and Environmental Sciences, University of Helsinki, 00014 Helsinki, Finland

Published online: 14 December 2023

advertising on digital platforms, affecting advertising effectiveness, user behavior and the revenue models of digital companies. Understanding these impacts is vital for both businesses and consumers. Furthermore, recommendation algorithms have behavioral influences, suggesting products, content or actions to users. These influences can be used for positive purposes but also for manipulation or harm, making it essential to study the mechanisms behind these behavioral effects and promote responsible technology use. Lastly, recommendation algorithms play a critical role in promoting or hindering fairness and equity by shaping the content users see. Biases in these algorithms can perpetuate societal inequalities, necessitating research and policymaking efforts to address algorithmic bias and ensure equitable access to opportunities and resources. Additionally, studying recommendation algorithms enhances transparency and accountability, shedding light on the often opaque and proprietary algorithms used by tech companies, allowing users to make more informed choices. Policymakers require a deep understanding of these algorithms to craft effective regulations that balance innovation with societal well-being, highlighting the importance of comprehensive research in this field.

1.2 Literature review

Recommendation systems use and process the digital footprint of millions of users mainly to retrieve and exploit their cognitive biases, and therefore, they can personalize their recommendations and deeply hit their users. For example, Youyou et al. (2015) demonstrated that Facebook likes can accurately predict users' sensitive information such as sexual orientation, political views, personality traits, use of addictive substances, personal separation and religious belief. In this case, they reported with 200–300 likes on Facebook, and the artificial intelligence behind the recommendation algorithm in this platform can predict users' personality traits better than their family members or even better than themselves. As algorithms become smarter and computationally more complex, they need even less input data to predict users' information. For example, Kosinski (2021) demonstrated that artificial intelligence can predict a person's political orientation only by their facial images with accuracy up to 72%, which was found to be higher than accuracy of judgments made by humans.

There are many societal implications with such smart and powerful algorithms. For example, recommendation systems can affect and control public participation in political affairs (Feezell et al. 2021). In this context, digital manipulation of democratic elections is a very important implication of recommendation systems in the realm of politics. Epstein and Robertson (2015) studied that voters' choices are significantly influenced by internet search rankings, mostly because they trust and favor higher-ranked results over those that are rated lower. In this study, researchers investigated whether search results could be manipulated to change the preferences of unsure voters in democratic elections given their apparent authority. According to their findings, biased search rankings can change the voting preferences of unsure voters from 20 to 80%, while a democratic election is won on average with a margin of 7%. Due to such serious and important impacts, digital companies have always been an attractive target for lobbying groups (Kreiss et al. 2018).

Based on such social implications, recommendation systems have a critical role in propagation of information in society, which can lead to indoctrination and mass customization of people. There are different underlying variables that affect the mechanism through which a message can propagate in a network. For example, the global structure of the network itself has been cited as a decisive factor shaping the diffusion of a message. In this case, Watts and Dodds (2007) used a series of computer simulations and interpersonal influence processes and they found that diffusion of a message over large cascades is not driven by influential people but by a large body of easily influenced individuals. Bakshy et al. (2011) further tested this hypothesis by investigating the attributes and relative influence of 1.6 million twitter users and tracking 74 million diffusion patterns that happened during a two-month time interval. They challenged the conventional assumption about celebrities' central role in the formation of public opinion. They found that only two variables of celebrities, namely number of followers and past influence, are significant to predict whether a celebrity's tweet can go viral, and the prediction accuracy was only 0.34, which is not a very good fit for a predictive model. They found that the density and structure of a network itself can properly explain whether there is possibility for a message to become viral. In other words, an intensively connected network among its nodes can properly predict the diffusion of a message no matter an influential or a regular person starts the cascade, and in a poorly connected network with many isolated nodes, there is no space for a message to propagate and again no matter who first starts it.

In addition to the global structure of the network, the content and attractiveness of the message itself is also an important factor affecting its successful spread across a network. Generally, those messages that can better stimulate and exploit users' cognitive biases have a better chance of propagation. For example, Vosoughi et al. (2018) modeled differential diffusion of true and false news through ~126,000 stories tweeted by ~3 million people more than 4.5 million times and they concluded that false news diffuses significantly *farther*, *faster*, *deeper* and *more broadly* than truth. They also found out that false stories are associated with fear, disgust and surprise in replies, and they showed a higher level of novelty compared to true news and people were more likely to share false news with such elements than true stories. In contrast, true stories were correlated with features such as anticipation, sadness, joy and trust. They also found that robots are indifferent in spreading false and true stories and the reason why false news spreads more than truth is because humans accelerate their diffusion and not robots. Like fake news, it seems conspiracy theories can also attract a fair amount of attention from users. In this case, Faddoul et al. (2020) presented a longitudinal analysis of YouTube's promotion of conspiracy videos. They found that during 2018–2019 and based on a monthly time interval, up to 10% of recommended videos were outrageously fake news or conspiracy theories. Interestingly, the name of YouTube's recommender engine, which drives 70% of YouTube views, is called Watch Time Optimization Algorithm and it acts like a "puppet master" over what users watch to keep their attention as much as it can by exploiting their cognitive biases.

Ethnic slurs can also trigger a chain of reactions from users, and therefore, they can accelerate propagation of a message or behavior in a network. In this case, Spörlein and Schlueter (2021) studied popular German political talk shows and they showed how ethnic insults propagate through network of comments in YouTube videos. They argue that ethnic insults imply social norms, they can trigger offensive behavior, and therefore, they can accelerate their contagiousness to other insulting commenting, and finally, it can lead to propagation of the behavior over the network. The spread of such comments is even faster and more intensive at the time of radical conflicts such as sexual assaults or terrorist attacks. Likewise, Andersson (2021) studied the relationship between impoliteness and homophily based on discussions underneath YouTube videos about Swedish climate activist, Greta Thunberg. She found out that recurrent impoliteness in the comment section of the videos about Greta Thunberg is not necessarily due to the design of social media that allows such behavior behind anonymity and social detachment, but rather this behavior can signal other like-minded users to form a front and build homophilous online communities.

In addition to network structure and enticing content, there are also a series of non-content variables such as duration, like count, positive comments, title, dislike count, view and negative comments that can predict the chances of a message or video going viral. In this regard, Halim et al. (2022) developed an AI-based framework to study non-content variables that affect the chances of a video to be landed in the trending category of YouTube.

There have been different solutions implemented to reduce the negative impacts of recommendation systems on society. For example, the design of a network itself can polarize or depolarize a context. In this case, Guilbeault et al. (2018) studied how political polarization can be solved between Democrats and Republicans about climate change. They studied whether it is the case that people not interacting with each other causes further polarization or it is the case that polarization is caused by putting people together. They designed an experiment in which they studied 2400 Republicans and Democrats and they found huge polarization between them in terms of predicting future climate change patterns. Then, they created various versions of a social media network where individuals from different political affiliations (Democrats and Republicans) had the option to either see or not see the political orientation of their counterparts. Their findings revealed that when people interacted without the ability to observe or anticipate the political views of the other party, it led to a significant increase in interaction and social learning. Furthermore, this depoliticized the context, suggesting that removing visibility of political affiliations can foster greater engagement and reduce polarization. Structural attributes of the network can also dictate depolarization to its users. For example, Santos et al. (2021) studied the effect of different social network typologies on the dynamics of polarization in social networks. They found that preferential link establishment between nodes with similar structures (i.e., nodes that share many neighbors between each other are structurally similar) encourages opinion polarization and connecting nodes with dissimilar structures can reduce polarization. In contrast to insulting, personal emotions were shown to be effective in bridging the gap between polarized groups. For instance, Kubin et al. (2021) studied YouTube comments about political topics such as abortion in Fox News and CNN channels and they found that personal experiences involving harm can bridge moral and political divides better than facts. In other words, people become more tolerant toward different narratives of truth when they are accomplished with subjective experiences involving harm than providing facts.

2 Research question and paper structure

Based on the literature review mentioned above, this paper aims to employ various modeling techniques to simulate the potential impacts of bias in YouTube's recommendation system on the diffusion patterns of climate change information in society. Accordingly, this study integrates several models, and these tools can function as a decision support system for the development and comparison of various information management scenarios. Put simply, the research question of this study is formulated as:

- How is climate change information diffused in society based on YouTube's recommendation algorithm and the political media landscape?

On this basis, in the second part of the paper, different methodologies such as web scraping, social network analysis and agent-based modeling are explained to, respectively, extract web data, quantify the bias in the recommendation algorithm and simulate the competition between different parties in pushing their own narrative of climate change in society. In the third section, we showcase and visualize the bias in YouTube's recommendation algorithm through social network analysis and statistical measures, and the results of agent-based modeling present the winner of the competition between different political parties in spreading climate change information from their own perspective. Ultimately, the paper concludes by discussing the main findings and suggesting future study directions.

3 Materials and methods

3.1 Public media bias ratings

There are different sources that rate media bias in news outlets. In this study, the bias rating result from (AllSides Media Bias Ratings 2022) has been adopted (Fig. 1), which has repeatedly been used as a source of data for media bias analysis in scientific papers. In this portal, the bias ratings are based on scientific and multi-partisan analysis. AllSides combines different crowdsourcing methods such as blind bias survey, editorial review, third-party analysis, independent review and community feedback for their ratings and in this study, their latest media bias results have been used (published in 2022). The media bias is categorized into five groups of Left, Lean Left, Center, Lean Right and Right. In this case, 67 news outlets were considered among which 66

Fig. 1 Media bias landscape based on multi-partisan and online political content. For detailed information on the rating method, please refer to source of the figure in www.allsides.com

had active YouTube channels (Alternet was the only outlet with no YouTube channel). Then, the most popular channels were selected from Left (including Lean Left), Center and Right (including Lean Right) parties to feed into web scraping and social network analysis for measuring the bias in YouTube's recommendation algorithm. The total number of subscribers in each category (Left, Center and Right) was also used as a basis to simulate the competition between different political perspectives in society for pushing their own narrative of climate change. The competition was simulated by growing a hypothetical society using an agent-based modeling approach.

3.2 Web scraping

In this study, the web data was extracted from YouTube using a visual web scraping tool offered by webscraper.io, which is a browser extension for Google Chrome. This extension uses a modular structure made of selectors and it navigates the scraper on how to traverse through the target website and what data to extract. The URL, title and the hosting channel of the videos were extracted. Accordingly, the data collection process for YouTube videos in this study involved several steps:

1) **Web data extraction tool**: The study utilized a visual web scraping tool provided by webscraper.io, which is a browser extension for Google Chrome (Baldis 2022). This tool allows users to extract data from websites in a structured manner.
2) **Data extraction**: The following details were extracted from YouTube for each video: the video's URL, title and the hosting channel.
3) **Preparation for unbiased data collection**: To ensure that the data collection process was as unbiased as possible, several steps were taken. Browsing history, cache data and cookies were cleared. This step helps to prevent any personalization effects caused by YouTube's recommendation algorithm, which could skew the results. The user signed out of the YouTube account, and the browser was launched in "incognito" mode. This further reduces the impact of algorithmic personalization.
4) **Initial search**: The homepage of each selected YouTube channel was loaded, and a search for videos related to the keyword "Climate Change" was conducted on each channel to retrieve videos exclusively from the respective channel at this stage.
5) **Featured videos**: The top 30 videos resulting from the initial search were selected. These videos were referred to as "featured videos" in the paper.
6) **First-level data collection**: For each of the featured videos from the previous step, additional data was collected by clicking on them. Specifically, the study

checked which videos and from which channels were recommended on the right panel of the webpage. This was referred to as "recommended videos level 1."

7) **Second- and third-level data collection**: The web scraper was configured to go even deeper into each recommended video in level 1, extracting video recommendations for two more levels ("recommended videos levels 2 and 3")
8) **Structured data file**: The result of this data collection process was a structured data file. This file allowed users to trace back three levels of recommended videos from the homepage of each media's YouTube channel.

The data collection process is designed to minimize bias and provide a structured dataset for subsequent analysis and modeling. Figure 2 shows the schematic representation of the network typology used in this study to scrap YouTube data.

3.3 Social network analysis

Social network analysis is a mathematical framework for studying social structures using networks and graph theory. The structure of a social network is characterized by nodes (i.e., videos and hosting channels) and links (i.e., hyperlink URLs to videos and channels) that connect nodes to each other. In this study, Gephi version 0.10 (Bastian et al. 2009) was used to study the network structure of recommended videos and their hosting channels. In this case, a directed network was used to measure the frequency and distribution pattern of recommended videos and their hosting channels based on web data derived from the last step and check whether there is a specific biased pattern fitting into political agenda of each channel. The descriptive statistics of distribution pattern of recommended videos and their hosting channels were then calculated, and the Q–Q plots were drawn to visualize the pattern of the distribution.

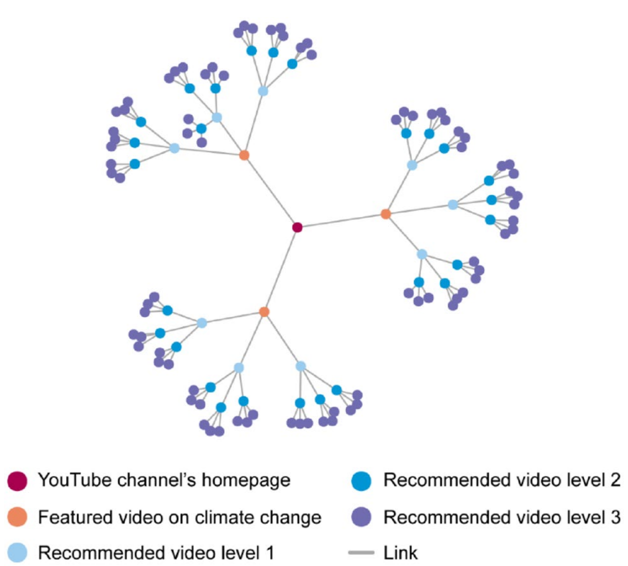

Fig. 2 Schematic representation of network typology used in this study to extract web data from YouTube and explore the bias in its recommendation system

3.4 Agent-based modeling

An agent-based model is a computational tool for simulating the behavior, actions and interactions of autonomous agents with a system of values that interact locally, and the result of their collective interactions can lead to emergence of a global pattern. Agent-based models combine different elements from game theory, complex systems, computational sociology and evolutionary programming. They have a wide variety of applications in biology, ecology and social science. In case of this study, this tool was used to simulate a theoretical society based on two-step flow of communication model (adopted from Hilbert 2023) to study the competition between Left, Center and Right media outlets in propagating climate change information in society from their own perspective. The relative size of the media for each party (Left, Center and Right) was configured based on the total number of subscribers to all their respective channels in YouTube. Three media outlets were introduced to the model as patch-based agents on which moving agents can walk onto and collect information. Moving agents represent humans in a society that can collect and store information in their memory whenever they interact with a media patch. The agents were programmed to wander through agent-based environment in a random pattern such that they take a 60-degree turn to a random direction during each model run. Put simply, every time an agent walks onto a media patch of a specific party, they consume one unit of information from that media and add it to their memory and then keep track of the same process through the next model runs. Based on the concept of two-step flow of communication (Baran and Davis 2014), the agents were able to communicate with each other and affect each other's performance.

According to two-step flow of communication model, most people are not directly influenced by mass media, and instead, their beliefs and opinions are shaped by opinion leaders who receive and interpret the message from the mass media and put them into context. In other words, opinion leaders or influentials are those who first get exposed to a specific media content and then interpret it according to their own opinion. The influentials then infiltrate the general public to indoctrinate the opinion into their followers. This process was internalized into the model through the interaction and information exchange between agents such that when two agents with information from different media outlets meet each other at the same point in the environment, the one who has accumulated less information from a media patch, sets its memory back to 0. In other words, it is influenced by a more influential agent, who has consumed more information from a different media outlet. Since the relative size of the media patches correspond to the number of subscribers from all their respective channels in YouTube, an additional element was added to the model to configure the results under different scenarios. In this regard, since the size of Left YouTube channels vastly dominate the media landscape, a series of weights were assigned to the message from the Right channels to study and compare the effect of media size and the weight of a message together on propagation of information in a society. In the context of this research, the weight of the message refers to the level of effort that a media outlet invests in crafting their message to make it more appealing, convincing and impactful for their target audience. In real-life terms, the weight of the message reflects the attractiveness and persuasive power of the content in capturing the attention of viewers and encouraging them to embrace the new message being conveyed. By assigning a weight to the message, we aimed to capture the potential influence and reach of the media's narrative, highlighting its ability to displace prior messages and effectively engage a wider audience. In other words, although Right media are smaller in size, what would happen if they allocated their resources to tailor their message in a very engaging and convincing manner compared to other media with bigger pool of subscribers? This question was explored by assigning weights to those agents carrying information from Right media in the environment. The set of weights was 100, 10,000 and 1,000,000, respectively. Put simply, agents from Left and Center media can beat Right agents in information exchange only if they have already accumulated 100, 10,000 and 1,000,000 times more information than Right media agents; otherwise, the Right agents will win and set the information value of their counterparts back to 0. By comparing the results under such scenarios, one can study whether it is the size of the media or the weight of the message that affects the success of a political party to push their own narrative of climate change in society. As a comparative basis to the weighted scenarios, one more scenario was added in which only agents of Left and Right media compete but the conditions for the size of their media and the weight of their messages were absolutely equal.

The model parameters included three counters, measuring the real-time count of messages from various parties being carried by agents. Additionally, a slider was used to set the initial population size at the beginning of the model run. An integral component was the dynamic graph, which tracked the total number of messages associated with each party as they were being processed. This dynamic graph was used to monitor and analyze the flow of messages throughout the model's operation, and the result of the graph was exported in the form of an excel file to draw Fig. 5. Netlogo version 6.3.0 (Tisue and Wilensky 2004) was used to configure the agent-based model.

Figure 3 demonstrates the overall flowchart of this study.

4 Results

4.1 Public media bias ratings

The number of subscribers for each media channel was extracted from YouTube (Table 1). Accordingly, Left and Lean Left channels together totally dominate the media landscape in terms of both the frequency of channels and the total number of subscribers. The total number of subscribers to Left, Center and Right channels was 80,965,200, 23,220,700 and 26,296,350, respectively. The most popular channels across Left, Center and Right showed less difference in terms of the size of subscribers. CNN, BBC News and Fox News were the most subscribed channels for Left, Center and Right parties with 14,600,000, 13,700,000 and 10,200,000 subscribers, respectively. In addition to having a smaller pool of subscribers, the drop rate in the number of subscribers across Center and Right channels is much bigger than that of the Left media, which means the majority of followers are only subscribed to first few channels. In other words, Left channels have a much bigger network of audiences from society, higher diversity in their media landscape and less inequality across their media outlets. CNN, BBC News and Fox News were selected for web scarping and analyzing the bias in YouTube's recommendation system, and the total number of subscribers was used as an input to configure the agent-based model.

4.2 Web scraping

CNN, BBC News and Fox News are not only impactful news organizations but also the most subscribed YouTube news channels among all other media outlets from their respective political parties. This distinction amplifies their influence on the digital landscape, where millions of viewers from around the world actively engage with their content and contribute to the ongoing discourse on climate change. With millions of viewers worldwide, these news outlets have the power to inform, educate and sway public opinion on climate-related issues. They provide a spectrum of perspectives, from progressive to conservative, which resembles the diversity of opinions on this matter.

The web scraping process for these channels was undertaken separately, and it resulted in structured data files such that a user could feed them into social network analysis with minor data wrangling efforts. The extracted data included the title and URLs of the recommended videos and their

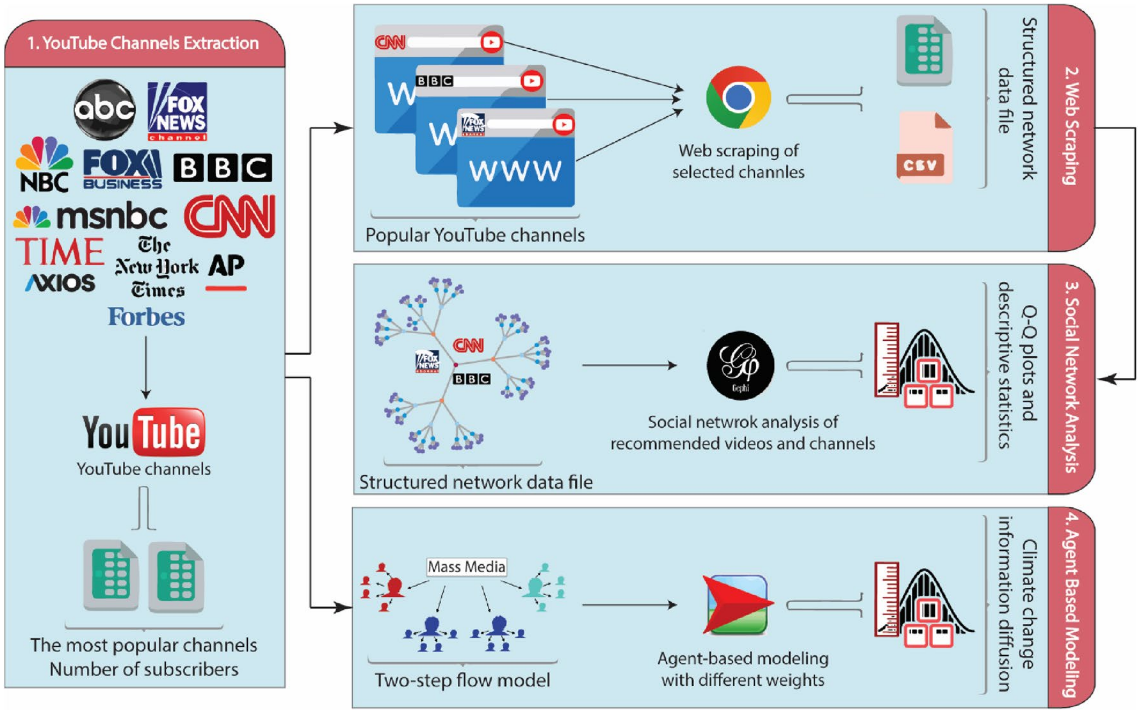

Fig. 3 Study flowchart

hosting channels. The data for recommended videos and channels was first divided and then fed into Gephi to separately analyze the bias in YouTube's recommendation algorithm for both recommended videos and YouTube channels.

4.3 Social network analysis

The results of social network analysis of YouTube's recommendation algorithm are presented in Table 2, 3 and 4. Table 2 shows the top 10 climate change-related recommended videos and YouTube channels to a Fox News viewer. As the titles of the videos indicate, the content of all these videos has a skeptical stance toward climate change. Fox News itself, by a big margin to the second-most suggested channel, is the most recommended channel to its own users, and most of the other top recommended channels are labeled under Right category in media landscape. Forbes Breaking News from Center and The Young Turks and MSNBC from the Left media are also top recommended channels to Fox News users but with a big difference to Fox News self-recommendations. The recommendations accounting for these top 10 videos and channels cover 15% and 31% of all recommendations for recommended videos and channels, respectively, which shows they have a very big impact on overall performance of YouTube's recommendation algorithm.

Table 3 shows top 10 recommended videos and channels to BBC News users. All top recommended videos are from BBC News itself and it seems the content of all videos acknowledge climate change as a global concern except one video which criticizes the concept of climate change in a political context (video title: Vladimir Putin criticizes Greta Thunberg's UN speech on climate change—BBC News). All top recommended channels are labeled as either Center or Left (only WION is labeled as slightly Right), which shows BBC News is more on the Left side when it comes to climate change and this matter is supported by YouTube's recommendation algorithm as well. The top 10 recommended videos and channels, respectively, cover 16% and 24% of all recommendations made to BBC News users, and like Fox News, BBC News self-recommendations to its channel totally dominate the performance of the recommendation algorithm.

Table 4 shows the top 10 recommended videos and channels related to climate change to CNN users. All the top recommended videos predominantly originate from CNN itself, while the highest recommended channels consist of TEDx Talks, StarTalk and CNN, with minimal variations between them. All recommended videos confirm climate change as a global concern (the second top recommended video is a debate between a democrat and republican on climate change) and top recommended channels are either Left or Center which has a similar stance toward climate change. These top 10 recommendations cover 16% and 27% of all suggestions made to CNN users.

It seems YouTube's recommendation algorithm is biased toward Left and Center media as they have a

Table 1 Media outlets with active YouTube channels and the size of their subscribers (retrieved 11 January 2023)

Left and lean left		Center		Right and lean right	
Channel's name	Subscribers size	Channel's name	Subscribers size	Channel's name	Subscribers size
CNN	14,600,000	BBC News	13,700,000	Fox News	10,200,000
ABC news	14,000,000	Wall Street Journal	3,970,000	Daily Wire	2,960,000
Vox	11,000,000	Reuters	1,830,000	Daily Mail	2,200,000
NBC news	7,350,000	The Hill	1,530,000	Fox Business	2,060,000
MSNBC	5,450,000	Forbes	1,350,000	Newsmax	2,050,000
CBS news	4,710,000	NewsNation	658,000	BlazeTV	1,640,000
The New York times	4,230,000	MarketWatch	70,300	One American News Network	1,410,000
Guardian news	2,960,000	Newsweek	62,400	CBN News	828,000
USA TODAY	2,350,000	Axios	33,300	ReasonTV	807,000
Washington Post	2,120,000	The Christian Science Monitor	15,300	New York Post	805,000
Associated press	2,040,000	RealClearPolitics	1,400	The Epoch Times	543,000
The nation	1,510,000			Breitbart News	350,000
Bloomberg markets and finance	1,440,000			Daily Caller	253,000
Democracy now!	1,270,000			Washington Free Beacon	62,200
Time	1,230,000			The Post Millennial	30,000
Buzzfeed news	806,000			IJR—Independent Journal Review	26,100
The new yorker	766,000			National Review	17,900
HuffPost	711,000			The Federalist	17,200
The Atlantic	647,000			The Washington Examiner	14,800
Insider news	454,000			The Washington Times	9750
NPR	410,000			The American Conservative	6200
The intercept	245,000			The Dispatch	3300
Slate	231,000			Wall Street Journal Opinion	1520
POLITICO	116,000			American Spectator	1380
Jacobin	110,000				
Mother jones	74,600				
ProPublica	62,000				
The daily beast	39,400				
Yahoo news	33,200				
Total number of subscribers	80,965,200		23,220,700		26,296,350

chance to be among top recommended channels to Fox News users, but the opposite is hardly true. Center media, TEDx Talks channel, appears to be among top recommended channels to all users of CNN, BBC News and Fox News followed by another Center media, DW Documentary, which is located among top recommended channels to CNN and BBC News users. In general, there are four YouTube channels commonly found among the top recommendations for both CNN and BBC News viewers. However, the overlap with Fox News is limited to only one channel. In other words, the network of Fox News users is less connected to the network of users from Left and Center, which means users of CNN and BBC News are less likely to end up in Right media, but Fox News users have a fair chance of getting recommendations from Center and Right channels.

Table 5 presents a series of global statistics for networks of recommended videos and channels to CNN, BBC News and Fox News users. For example, all average values are greater than medians for recommended videos and channels and it means the data is heavily skewed to the right and the skewness values also confirm the same issue. (Skewness > 0.5 means data distribution is not symmetric.) This issue is also visually presented in Fig. 4. In other words, the distribution patterns of links to recommended videos and channels suggest that an extremely small number of videos and channels receive a disproportionately high number of recommendations, while a vast number of unique videos

Table 2 Top 10 most recommended videos and YouTube channels to Fox News users

Video title	Recommendation frequency
Berkeley's climate change solution: population control	6357
Murray: Climate activists are moving 'too far too fast'	3584
Tucker: CNN's climate change town hall was an act of wanton cruelty	3340
Gutfeld: The Left's hypocritical world of climate change	3323
How Patagonia took it upon itself to fight climate change	3145
Climate protest backfires as parolee can't get to work	3041
What's the biggest threat—China, Russia or climate change? | Americans Weigh In	2164
We need more Republicans to say climate change is real: Jessica Tarlov	2121
Tucker roasts Jane Fonda for claiming racism causes climate change #shorts	2079
McEnany: Climate change is now to blame for everything	1844
Sum of the links made through these top 10 recommendations	30,998 (15% of all links)
Total number of links made in the whole dataset	213,202
Channel name	**Recommendation frequency**
Fox News	18,126
Forbes Breaking News	7970
Sky News Australia	5602
The Young Turks	3517
TEDx Talks	3007
Jordan B Peterson	2676
John Stossel	2671
The Telegraph	2309
Fox Business	2240
MSNBC	1918
Sum of the links made through these top 10 recommendations	50,036 (31% of all links)
Total number of links made in the whole dataset	159,921

and channels receive very few recommendations through YouTube's recommendation algorithm. This biased distribution indicates a preference for specific content within the algorithm.

4.4 Agent-based modeling

The results of agent-based modeling based on two-step flow of communication model are presented in Fig. 5. The results of agent-based modeling were exported in the form of excel files, and the graphs were visualized in R environment. This figure visualizes the results of the competition between the media from different political parties in spreading climate change information in society from their own perspective. Figure 5a shows the result of competition between only Left and Right media outlets under equal conditions. In other words, equal number of subscribers was assumed for both media with equal weights to their corresponding message. The figure clearly illustrates a cyclical pattern in which Left-leaning media sources and Right-leaning media sources take turns in dominating the information landscape. Initially, Fig. 5a portrays a scenario where Right-leaning media outlets wield a significant influence over public discourse. However, as time progresses, the Left-leaning media outlets begin to gain momentum, eventually surpassing their Right-leaning counterparts in terms of the volume and impact of their messaging, and the same cyclic pattern is constantly repeated over time. In essence, according to the figure, there is a recurrent pattern where Left-leaning and Right-leaning media sources take turns in dominating the societal conversation on climate change and there is no single absolute winner and both sides have equal chance in preaching their ecological agenda. In Figs. 5b–d, the model is configured for the size of the media according to the total number of subscribers from all their respective channels in YouTube and the competition was simulated with higher weights assigned to the message of Right media. The weights were 100, 10,000 and 1,000,000 for Figs. 5b, c and d, respectively. Accordingly, the Left media under all scenarios dominate Right and Center media in spreading their own narrative of climate change in society. In other words, the size of the audience has a bigger impact on the success of a respective media than the weight of its message. If a media has a smaller network of subscribers, their message will not reach

Table 3 Top 10 most recommended videos and YouTube channels to BBC News users

Video title	Recommendation frequency
Arnold Schwarzenegger calls leaders 'liars' over climate change—BBC News	4573
Global water crisis looming, UN says—BBC News	3872
Climate Change: Professor Brian Cox clashes with sceptic Malcolm Roberts—BBC News	3703
Why should we care about climate change? BBC News	2986
Vladimir Putin criticises Greta Thunberg's UN speech on climate change—BBC News	2627
Climate Change: Your carbon footprint explained—BBC News	2404
Is climate change to blame for extreme heat in Europe?—BBC News	2355
Arctic polar bears "face near-extinction within decades" warn scientists—BBC News	2321
UN scientists say it's 'now or never' to limit global warming—BBC News	1649
Global warming set to break key 1.5Ctemperature limit for first time #Shorts #ClimateChange#ElNino	1532
Sum of the links made through these top 10 recommendations	28,022 (16% of all links)
Total number of links made in the whole dataset	171,534
Channel name	**Recommendation frequency**
BBC News	7993
BBC Learning English	4418
TEDx Talks	3634
Sky News	2631
WION	2379
DW Documentary	2334
PBS NewsHour	2290
NBC News	2060
CNN	1867
Channel 4 News	1695
Sum of the links made through these top 10 recommendations	31,301 (24% of all links)
Total number of links made in the whole dataset	128,676

beyond a certain limit no matter how big the weight of their message is, and in contrast, if a media has a bigger pool of followers, their message will reach to broader network of people in society no matter how small the weight of their message is. Accordingly, owing to their relatively equal network sizes, the graphs for Right and Center media outlets illustrate similar patterns (Figs. 5b–d), with competition between these two parties being more balanced compared to the Left media outlets, which completely dominate the information landscape regarding climate change discourse.

5 Discussion

It is estimated that over 2 billion people worldwide collectively spend 1 billion hours watching YouTube videos each day, which positions YouTube as the second-most visited website globally, right after Google. Moreover, it holds the distinction of being the most popular video-sharing social network on the planet. YouTube's recommendation algorithm is the main mediator that decides which videos to be suggested to such a huge number of users from a database with millions of video records. It is estimated that 70% of total watch time in the entire platform is driven by its recommendation engine (Covington et al. 2016), and therefore, its massive social and ethical implications are undeniable. Understanding the bias and the mechanism through which YouTube's recommendation system works can bring transparency to both content creators and users (Wojcicki 2021), and it also helps decision makers in the realm of digital companies to regulate and control the possible implications of such technologies through proper legislative channels (Brossard 2013; Iyengar and Massey 2019). In this case, Democrats introduced Protecting Americans from Dangerous Algorithms Act. In 2021 to hold YouTube accountable for the social effects of its recommendation system (Tutt 2017).

Biased recommendation systems can actively change and form information landscape in a society. In the case of climate change, recommendation systems have a major impact on formation of attitudes in society toward the issue and climate change itself has fundamental implications in economic and political domains (Törnberg 2022). In other words, users only following conservative and Right media are more likely the disagree with the effects of climate

Table 4 Top 10 most recommended videos and YouTube channels to CNN users

Video title	Recommendation frequency
Climate change is affecting the food you eat. Here's how	5023
Bill Nye, Marc Morano on Global Warming	3156
Is climate change to blame for this?	2771
Undeniable climate change facts	2359
DeGrasse Tyson: We have to believe science on climate change	2354
Climate change made these women question whether or not to have children	2306
OBAMA CLIMATE CHANGE-SEA LEVEL IN NY	2179
Obama slams GOP on climate change	2098
Scientist laughs at climate change skeptics	2003
Clinton: Trump called climate change a Chinese hoax	1999
Sum of the links made through these top 10 recommendations	26,248 (16% of all links)
Total number of links made in the whole dataset	165,206
Channel name	**Recommendation frequency**
TEDx Talks	5822
StarTalk	5639
CNN	5311
TED	3751
PBS NewsHour	2826
CBS News	2157
Vox	1985
The Obama White House	1960
DW Documentary	1851
Wall Street Journal	1842
Sum of the links made through these top 10 recommendations	33,144 (27% of all links)
Total number of links made in the whole dataset	123,912

Table 5 Descriptive statistics derived from distribution of links to recommended videos and channels based on YouTube recommendation system (in this table REC means recommended)

YouTube channel	Fox News		BBC News		CNN	
	REC videos	REC channels	REC videos	REC channels	REC videos	REC channels
Number of unique REC videos/channels	21,871	4360	18,008	4190	16,748	4403
Average number of links to a REC video/channel	9.75	36.68	9.53	30.71	9.86	28.14
Median of all links	1.00	1.00	1.00	1.00	1.00	1.00
Mode of all links	1.00	1.00	1.00	1.00	1.00	1.00
Skewness of links distribution across all REC videos/channels	38.46	36.32	34.45	20.15	32.66	17.95
Kurtosis of links distribution across all REC videos/channels	2,062.99	1716.55	1,498.40	610.91	1,458.47	418.92

change and followers of Left and Center media are more likely to consider climate change as a serious issue that requires international collaboration and global governance. Put simply, such technological systems can turn an ecological concern into a political agenda and therefore politicize the context and cause polarization in society.

The findings of this study suggest that YouTube's recommendation system is extremely biased toward certain videos and channels depending on which channel a user starts browning climate change content and it seems the content and political affiliation of highly recommended videos and channels fit the corresponding political affiliation of

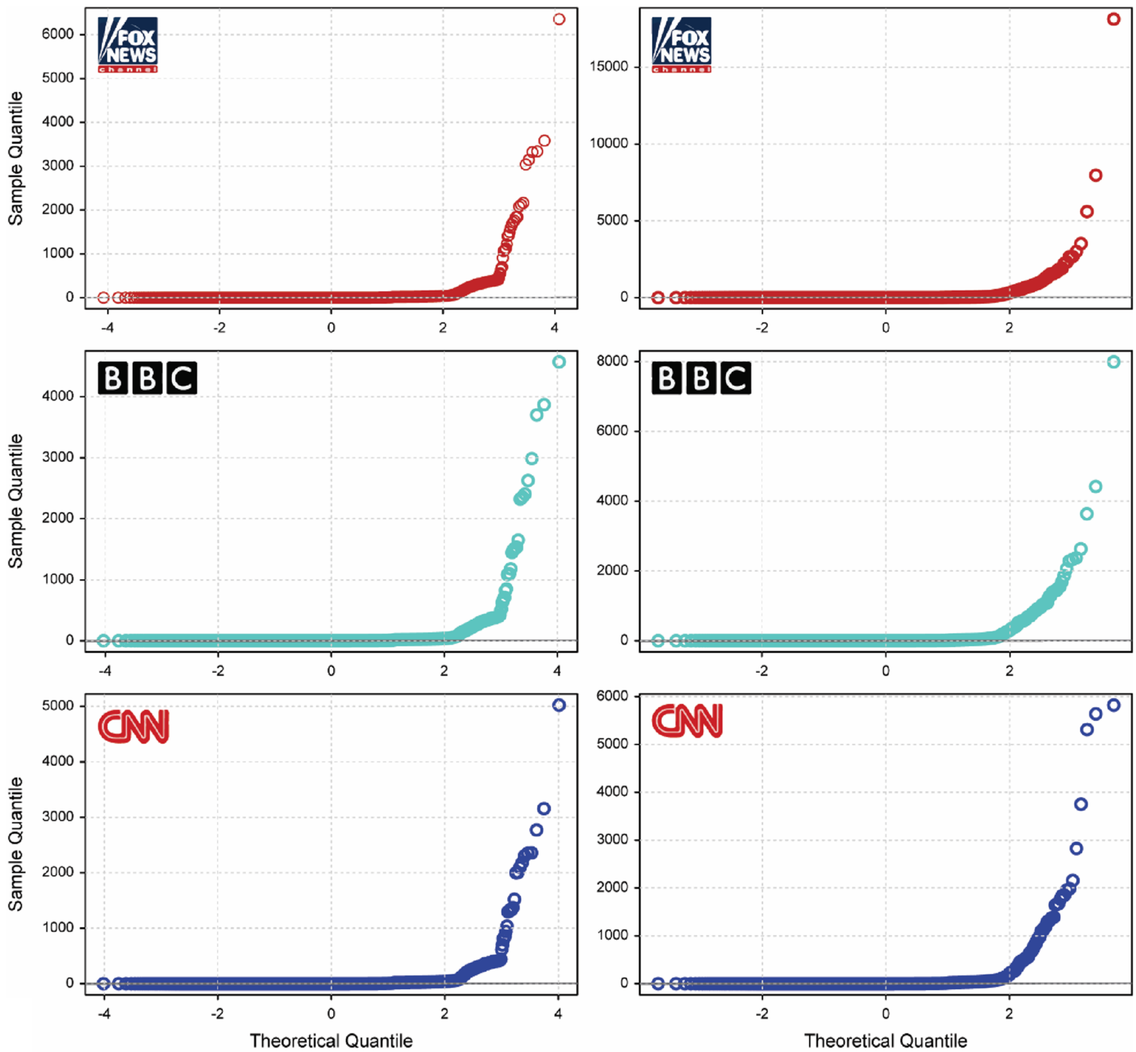

Fig. 4 Q–Q plots visualizing distribution patterns of link frequencies to recommended videos, (*left panel*) and channels (*right panel*) through YouTube's recommendation system

a specific channel. Therefore, future studies can study the social effect of such highly biased recommendation pattern in a more explicit way. For example, the comment section of these videos can provide a rich pool of data for emotions and attitudes toward climate change, and analyzing these comments across videos and channels can shed lights on the mechanism through which recommendation system causes polarization over climate change and related topics. YouTube provides valuable information in many different contexts that can be implemented in various branches of science (Jagiello et al. 2019; Kikuch et al. 2022; Gorgy et al. 2022; Sütcüoğlu et al. 2022; Silek and Topcuoglu 2023). Accordingly, the results suggest that an echo chamber exists in YouTube's recommendation algorithm. In discussions about news media, an echo chamber is a situation in which people are put in a closed and insulated system to only engage with beliefs or opinions that confirm their own, and therefore, the existing views are constantly reinforced, and opposition views are actively excluded from engagement. In other words, echo chambers exploit humans' cognitive biases such as confirmation bias to nudge the users into an effortless and self-amplifying loop to absorb their attention as much as possible and maximize their activity on the platform by pulling them deeper inside the matrix. Likewise, Kirdemir et al. (2021) found structural and systematic bias in YouTube's video recommendations with varying degrees between different experiments and the results of the study by Alfano et al. (2021) confirm the radicalization effect of YouTube's recommendation system on its users. With the same spirit, Markmann and Grimme (2021) investigated the polarizing effect of YouTube's recommendation algorithm with particular focus on "autoplay" function. They studied different recommendation pathways with personalized accounts and anonymous users, and they found that although

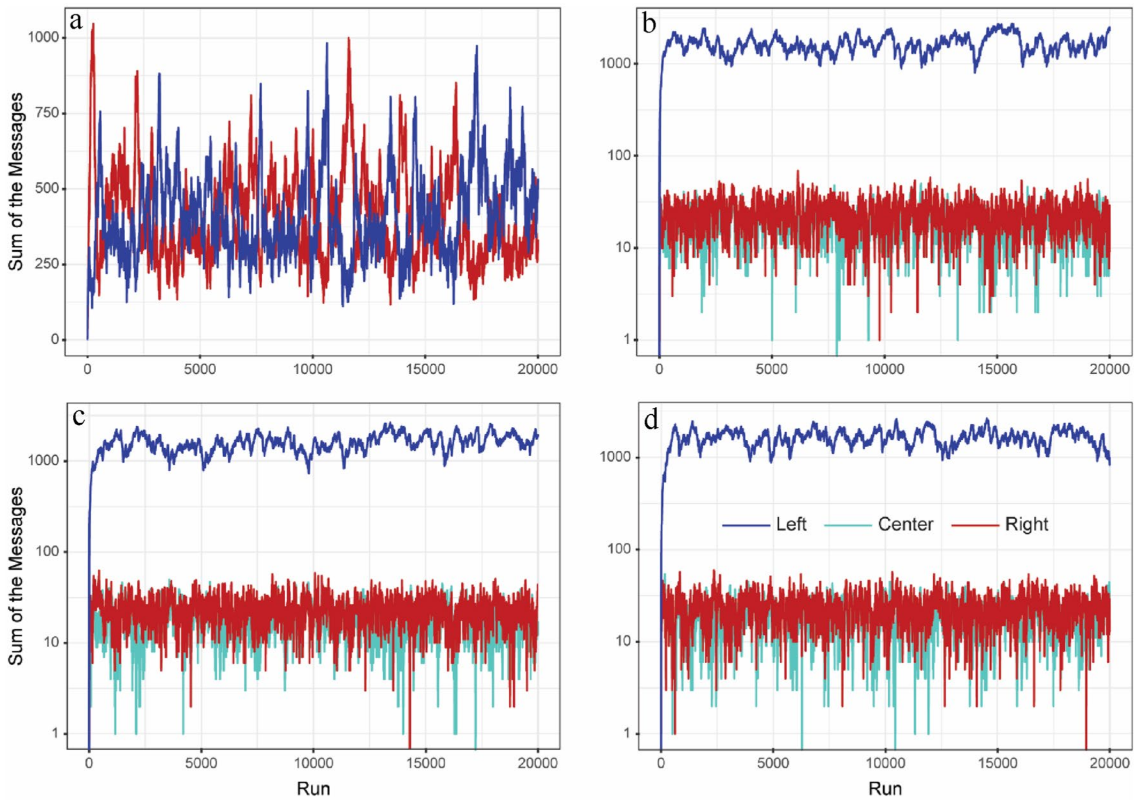

Fig. 5 Results of agent-based modeling based on two-step flow of communication model. This figure visualizes the competition between mass media from Left, Center and Right political parties in propagating climate change information in society from their own perspective. (**a**) This figure serves as a comparative basis to show under equal conditions (the same weight for the message and the same audience size), there is an equal chance for each media to propagate their message and it is reflected in the recurrent cyclic pattern through which Left and Right parties switch turns over time in dominating the media landscape. (**b–d**) These figures visualize the competition between Left, Center and Right media in distributing climate change information based on the size of their subscribers in YouTube and unequal weights for their message. The weights were only assigned to Right media to increase the power of their message in competing with Center and Left media, who have a bigger number of subscribers together than the Right channels in YouTube. The weights are 100, 10,000 and 1,000,000 for b, c and d, respectively. In b, c and d figures, the y axis is log-transformed

YouTube uses such mechanism to capture users' attention, the increasing public awareness regarding the effect of recommendation systems has led to algorithmic parametrization of video recommendations to moderate content. In the same way, Hosseinmardi et al. (2021) reported similar results about the effect of YouTube's recommendation system that is dominated by mainstream and largely centrist sources compared to far-right and anti-woke groups. For example, they did not find any systematic correlation between anti-woke and far-right channels in this platform.

Crane and Sornette (2008) used a time series of daily views for almost 5 million YouTube videos to model and predict the reaction of a social system to endogenous and exogenous variables. In other words, by monitoring the reaction of social systems and collective human dynamics in response to YouTube videos over time, we can model the relationships between the digital content and actual humans' reaction to these content in the real world. Based on such models, it is possible to predict what type of content can stimulate a specific response from societies. This concept is operationalized in, for example, the GDELT project (The GDELT Project 2022). The GDELT project monitors the world's broadcast print and web news in every country and in more than 100 languages and detects and maps peaceful or aggressive reactions to such news. By monitoring and collecting such data over time and projecting the corresponding response from social systems in a spatial manner, this project can model the spatiotemporal dynamics of the relationships between news and the respective response to them on the ground. Accordingly, developers can use such models to stimulate a target society into a specific response by broadcasting engineered news into their mass media.

The results of the agent-based model confirm the findings of Watts and Dodds (2007) and Bakshy et al. (2011). Put simply, the size and structure of the network are decisive factors on successful spread of a message in the society. If the network is already small and isolated, for example Right YouTube channels have a smaller pool of subscribers and they are actively excluded from being recommended to Left and Center media, the message cannot spread beyond a certain limit no matter what is the weight of the message or how influential the preacher of the message is. Santos et al. (2021) reported a similar concept using network terminology as preferential network typology increases polarization and differential link establishment between networks with different structures can reduce polarization. However, when comparing the messages within the same network, those messages that are more enticing and engaging have a higher chance of diffusion in that particular network no matter whether a celebrity or a regular person starts the cascade. Therefore, it is important to differentiate between diffusion of different a message across various networks or within the same network as driving forces of information propagation in these networks are different.

It is important to note that not only the recommendation system, but also the design of YouTube causes radicalization and encourages partisan sorting among its users. In this regard, Guilbeault et al. (2018) reported the same findings when studying the attitudes toward climate change between Democrats and Republicans. It is because YouTube users can easily spot the logos and names of channels, and therefore, their cognitive biases will prevent them from clicking on videos of the opposing groups even if those videos get a small chance to pop up into their recommendation list. Abul-Fottouh et al. (2020) and Andersson (2021) reported the same results as people tend to be attracted to like-minded individuals and groups in social media.

The influence of various factors on YouTube's recommendation algorithm, such as watch time, likes and shares, represents an intriguing area for further research. Exploring how these factors affect the dissemination of climate change information within the platform's recommendation system could provide valuable insights into the reliability and applicability of findings of this paper. Future studies could focus on investigating the specific impact of these factors on the recommended content and the subsequent diffusion of climate change narratives. By incorporating these parameters into agent-based models, researchers can aim to create more realistic simulations that capture the complex interplay between political media landscapes, recommendation systems and the spread of climate-related information. Additionally, considering the role of user engagement metrics, social network dynamics and demographic factors in shaping YouTube's recommendations could contribute to a more comprehensive understanding of the underlying mechanisms at play. This expanded investigation would not only shed further light on the current patterns but also inform strategies for improving the accuracy and transparency of recommendation algorithms in addressing climate change discourse.

6 Conclusions

The results of this study provide evidence that YouTube's recommendation system is highly biased in recommending certain videos and channels to users of different channels with a political stance toward climate change. In other words, there is a tight correlation between ecological concerns of a certain party and their political agenda and YouTube's recommendation algorithm can potentially reinforce this relationship by recommending content that fits the political taste of subscribers to these channels.

The result of the competition between different political views toward climate change was simulated based on the size of subscribers for each party in YouTube, and results suggest Left media that are already controlling a sizable chunk of the network in YouTube will dominate the information landscape by pushing their own narrative of climate change. In addition, Center media have the same stance toward climate change as Left news channels, and they can amplify the spread of their respective message as there is more connectivity between their media than with media from the Right. Future studies can further explore and model the relationships between the size, the weight and the number of media to produce additional insights on the outcome of competition between different narratives in society. Such findings shed light on the mechanism through which public perception on climate change is created and shaped by recommendation systems and digital companies.

Acknowledgements The author would like to express sincere gratitude to Aruna Rani Deb for her kind assistance in collecting the relevant literature. Additionally, the author extends appreciation to anonymous reviewers whose constructive comments contributed to the enhancement of this paper.

Author contributions The development of conceptual framework, data collection and analysis and visualizations have been done by the corresponding author of the paper Y.S.

Funding Open Access funding provided by University of Helsinki (including Helsinki University Central Hospital).

Declarations

Competing interests The authors declare no competing interests.

Conflict of interest None.

References

Abul-Fottouh D, Song MY, Gruzd A (2020) Examining algorithmic biases in YouTube's recommendations of vaccine videos. Int J Med Informatics 140:104175

Alfano M, Fard AE, Carter JA, Clutton P, Klein C (2021) Technologically scaffolded atypical cognition: The case of YouTube's recommender system. Synthese 199(1):835–858

AllSides Media Bias Ratings. (2022). Media bias rating methods | How AllSides rates media bias, [online] Available: https://www.allsides.com/media-bias/media-bias-rating-methods.

Andersson M (2021) The climate of climate change: impoliteness as a hallmark of homophily in youtube comment threads on greta thunberg's environmental activism. J Pragmat 178:93–107

Bakshy E, Hofman JM, Mason WA, Watts DJ (2011) Everyone's an influencer: quantifying influence on twitter. In: proceedings of the fourth ACM international conference on Web search and data mining (pp. 65–74).

Baldis M (2022) Making web data extraction easy and accessible for everyone. https://webscraper.io/

Baran SJ, Davis DK (2014) Theories of mass communication: glossary. Introduction to Mass Communication. McGraw-Hill Education, New York City

Bastian M, Heymann S, Jacomy M (2009) Gephi: an open source software for exploring and manipulating networks. In: proceedings of the international AAAI conference on web and social media vol. 3 no. 1 pp 361–362

Brossard D (2013) New media landscapes and the science information consumer. Proc Natl Acad Sci 110(3):14096–14101

Covington P, Adams J, Sargin E (2016) Deep neural networks for YouTube recommendations. In: proceedings of the 10th ACM conference on recommender systems pp 191–198

Epstein R, Robertson RE (2015) The search engine manipulation effect (SEME) and its possible impact on the outcomes of elections. Proc Natl Acad Sci 112(33):E4512–E4521

Faddoul M, Chaslot G, Farid H (2020) A longitudinal analysis of YouTube's promotion of conspiracy videos. arXiv preprint arXiv:2003.03318.

Feezell JT, Wagner JK, Conroy M (2021) Exploring the effects of algorithm-driven news sources on political behavior and polarization. Comput Hum Behav 116:106626

Gorgy A, El Hawary H, Galli R, MacDonald M, Barone N, Thibaudeau S (2022) Evaluating the educational quality of surgical youtube® videos: a systematic review. Health Sci Rev 11:100067

Guilbeault D, Becker J, Centola D (2018) Social learning and partisan bias in the interpretation of climate trends. Proc Natl Acad Sci 115(39):9714–9719

Halim Z, Hussain S, Ali RH (2022) Identifying content unaware features influencing popularity of videos on YouTube: a study based on seven regions. Expert Syst Appl 206:117836

Hilbert, M. (2023). Computational Science Specialization, Coursera. https://www.coursera.org/specializations/computational-social-science-ucdavis

Hosseinmardi H, Ghasemian A, Clauset A, Mobius M, Rothschild DM, Watts DJ (2021) Examining the consumption of radical content on YouTube. Proc Natl Acad Sci 118(32):e2101967118

Iyengar S, Massey DS (2019) Scientific communication in a post-truth society. Proc Natl Acad Sci 116(16):7656–7661

Jagiello ZA, Dyderski MK, Dylewski Ł (2019) What can we learn about the behaviour of red and grey squirrels from YouTube? Eco Inform 51:52–60

Kikuchi Y, Nishimura I, Sasaki T (2022) Wild birds in YouTube videos: Presence of specific species contributes to increased views. Eco Inform 71:101767

Kirdemir B, Kready J, Mead E, Hussain MN, Agarwal N, Adjeroh D (2021) Assessing bias in YouTube's video recommendation algorithm in a cross-lingual and cross-topical context. In: international conference on social computing, behavioral-cultural modeling and prediction and behavior representation in modeling and simulation. Springer, Cham. pp 71–80

Kosinski M (2021) Facial recognition technology can expose political orientation from naturalistic facial images. Sci Rep 11(1):1–7

Kreiss D, McGregor SC (2018) Technology firms shape political communication: the work of microsoft, facebook, twitter, and google with campaigns during the 2016 US presidential cycle. Polit Commun 35(2):155–177

Kubin E, Puryear C, Schein C, Gray K (2021) Personal experiences bridge moral and political divides better than facts. Proc Natl Acad Sci 118(6):e2008389118

Markmann S, Grimme C (2021) Is youtube still a radicalizer? An exploratory study on autoplay and recommendation. Multidisciplinary International Symposium on Disinformation in Open Online Media. Springer, Cham, pp 50–65

Milano S, Taddeo M, Floridi L (2020) Recommender systems and their ethical challenges. AI & Soc 35(4):957–967

Milano S, Taddeo M, Floridi L (2021) Ethical aspects of multi-stakeholder recommendation systems. Inf Soc 37(1):35–45

Santos FP, Lelkes Y, Levin SA (2021) Link recommendation algorithms and dynamics of polarization in online social networks. Proc Natl Acad Sci 118(50):e2102141118

Silek H, Topcuoglu OB (2023) Analysis of YouTube videos as a source of information for reliability and effectiveness of cannabidiol oil in treatment of epilepsy. Epilepsy Behav 138:109017

Spörlein C, Schlueter E (2021) Ethnic insults in youtube comments: social contagion and selection effects during the German "refugee crisis." Eur Sociol Rev 37(3):411–428

Sütcüoğlu O, Özay Zİ, Özet A, Yazıcı O, Özdemir N (2022) Evaluation of scientific reliability and quality of youtube videos on cancer and nutrition. Nutrition 108:111933

The GDELT project. (2022). https://www.gdeltproject.org/

Tisue S, Wilensky U (2004) Netlogo: a simple environment for modeling complexity. Int Conf Comp Syst 21:16–21

Törnberg P (2022) How digital media drive affective polarization through partisan sorting. Proc Natl Acad Sci 119(42):e2207159119

Tutt A (2017) An FDA for algorithms. Admin l Rev 69:83

Vosoughi S, Roy D, Aral S (2018) The spread of true and false news online. Science 359(6380):1146–1151

Watts DJ, Dodds PS (2007) Influentials, networks, and public opinion formation. J Consumer Res 34(4):441–458

Wojcicki S (2021) Letter from Susan: our 2021 priorities. YouTube Off Blog 26:30–37

Youyou W, Kosinski M, Stillwell D (2015) Computer-based personality judgments are more accurate than those made by humans. Proc Natl Acad Sci 112(4):1036–1040